**3.1 随机变量和离散分布** 2020年1月19日09点34分

**例题3.1.1** 抛硬币 抛10次公平硬币,将正面出现的次数定义为随机变量,是样本空间中元素的变量.

**定义3.1.1随机变量** 设是某项实验的样本空间.定义在上的实数函数被称为随机变量.

**例题3.1.2** 测量人体身高 这里身高就是随机变量

**例题3.1.3** 资源需求 举例随机变量,这里对随机变量的定义稍微复杂

**随机变量的分布**

当某项实验的样本空间的概率已经测得,那么我们就可以确定每一个随机变量可能值的概率.设是实数子集使得是一个事件,设表示的值属于子集的概率.则等于实验结果使得的概率.用符号表示为

**定义3.1.2 分布** 随机变量的分布是形如所有概率的集合,其中是任意实数子集并使得是一个事件.

**例题3.1.4** 抛硬币一般化

**例题3.1.3** 资源需求一般化

**定义3.1.3 离散分布/随机变量** 如果随机变量只能取个不同值或至多无限个不同值.则我们称有离散分布或是离散随机变量.

**定义3.1.4 概率函数/p.f./支持** 如果随机变量有离散分布,则的概率函数(简称p.f.)定义为函数使得对于每一个实数,

闭集被称为的支持(分布).

注意:概率函数有时在其它教材被称为概率质量函数,或p.m.f.但是在本书中我们不使用此术语.

**例题3.1.6** 资源需求 对**例题3.1.3** 提出的随机变量的解答.

**例题3.1.7** 抛硬币

**定理3.1.1** 设随机变量的p.f.为.如果不是的可能值,则.同样,如果序列包含所有可能值,则.

**定理3.1.2** 设有离散分布,实线的每一个子集的概率可以从下列关系确定

**定义3.1.5 伯努利分布/随机变量** 随机变量只能取两个值0和1,且,则被称为参数为的伯努利随机变量.

**例题3.1.7** 抽奖号码

**定义3.1.6** 整数的均匀分布 设是整数.假设随机变量的值等可能地出现在区间上.则我们称为在区间上有均匀分布.

**定理3.1.3** 如果在区间上有均匀分布,则的p.f.是

**例题3.1.9** 零件损坏 假设一台机器生产的零件损坏率为,并假设这台机器生产了个零件,设随机变量是零件损坏的个数,显然的取值范围是区间.则的p.f.为

**定义3.1.7 二项分布/随机变量** 由公式所表达的离散分布被称为*参数是和的二项分布*.符合该分布的随机变量被称为*参数是和的二项随机变量*.

**3.2 连续分布** 2020年1月19日18点14分

**例题3.2.1** 资源需求

**定义3.2.1 连续分布/随机变量** 如果存在一个定义在实数上的函数,对于实数任意区间(开区间或闭区间均可),使得随机变量在该区间上取值的概率等于在该区间的积分,则我们称有连续分布或是连续随机变量.

**定义3.2.2 概率密度函数/p.d.f./支持** 如果有连续分布,则定义3.2.1所描述的函数被称为的概率密度函数(简称p.d.f.).闭集被称为的支持(分布).

根据上述定义,对于任意闭区间

每一个p.d.f.必须满足下列两个需求:

和

**注意:连续分布为单个值赋予的概率为0**.公式与和是相等的.因此,从连续分布的定义得到,如果有连续分布,对于每一个值,.但是并不意味着不会发生.

**例题3.2.2** 温度预报 引出下列定义

**定义3.2.3 区间上的均匀分布** 设和是给定的两个实数并且.设是随机变量且已知,对于的每一个子区间,属于给定区间的概率与该区间的长度成正比.我们则称随机变量*在区间上有均匀分布*.

**定理3.2.1** 设在区间上有均匀分布,则的p.d.f.为

**注意**:**密度不等于概率**.p.d.f.是无界限的;在处的p.d.f.无法给出在附近的概率,只能通过p.d.f.在附近的积分得到;由于无界区间的长度是无穷的,因此无法在无界区间定义均匀分布;p.d.f.可以定义在开区间和半开区间,并且定义相同,例如公式定义在上与定义在上是相同的.

**例题3.2.3** 不完整的p.d.f.

**注意:计算归一化常(系)数**.某些随机变量是无界的，如果仅仅计算有界范围内的概率分布，则p.d.f.需要除有界范围内的总概率.

**例题3.2.4** 计算p.d.f.

**例题3.2.5** 无界随机变量

**例题3.2.6** 无界p.d.f.

**例题3.2.7** 混合p.d.f. 对**例题3.2.5**的扩展

**3.3 累积分布函数** 2020年2月5日14点17分

**例题3.3.1** 电压 是**例题3.2.5**的延续

**定义3.3.1** **累积分布函数** 随机变量的分布函数或累积分布函数(缩写为c.d.f.)定义为

**例题3.3.2** 伯努利c.d.f.设为参数为的伯努利随机变量,即.则的c.d.f.为

**属性3.3.1** **非递减** 随着增大是非递减的;即,如果,则.

**属性3.3.2** 且.

**属性3.3.3** 右侧连续 如果满足任意值,则c.d.f.总是在右侧连续.

**例题3.3.3** 电压 是**例题3.3.1**的延续

**定理3.3.1** 对每一个值,

**定理3.3.2** 对所有的值和值使得,

**定理3.3.3** 对每一个值,

**定理3.3.4** 对每一个值,

**定理3.3.5** 设是连续分布,并且设和分别表示其p.d.f.和c.d.f.则在每一个处都连续,

并且

在所有上使得连续.

**例题3.3.4** 从c.d.f.计算p.d.f.

**例题3.3.5** 预测降雨量,引出分位数函数

**定义3.3.2 分位数/百分位数** 设是一个随机变量,其c.d.f.为.对于每一个严格位于0和1之间的值,定义是使得的最小值.则被称为**分位数[quantile]**或的100**百分位数[percentile]**.函数定义在开区间上被称为的**分位数函数[quantile function]**.

**例题3.3.6** 分位数函数的应用

**连续分布的分位数** 当随机变量的c.d.f.是连续的并且在的所有可能值的集合中一对一,则的反函数存在并且等于的分位数函数.

**例题3.3.7** 投资风险 分位数函数的应用,**没太认真看**

**例题3.3.8** 均匀分布的分位函数

**例题3.3.9** 二项分布的分位数

**定义3.3.3 中位数/分位数** 分布的1/2分位数或第50个百分位数称为其中位数.1/4分位数或25个百分点是下四分位数.3/4分位数或第75个百分位数称为上四分位数.